



# MHC Class II haplotypes of Colombian Amerindian tribes

## Citation

Yunis, Juan J., Edmond J. Yunis, and Emilio Yunis. 2013. "MHC Class II haplotypes of Colombian Amerindian tribes." *Genetics and Molecular Biology* 36 (2): 158-166. doi:10.1590/S1415-47572013005000014. <http://dx.doi.org/10.1590/S1415-47572013005000014>.

## Published Version

doi:10.1590/S1415-47572013005000014

## Permanent link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:11717600>

## Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA>

## Share Your Story

The Harvard community has made this article openly available.  
Please share how this access benefits you. [Submit a story](#).

[Accessibility](#)



## MHC Class II haplotypes of Colombian Amerindian tribes

Juan J. Yunis<sup>1,2,3</sup>, Edmond J. Yunis<sup>4</sup> and Emilio Yunis<sup>3</sup>

<sup>1</sup>*Departamento de Patología, Facultad de Medicina e Instituto de Genética, Universidad Nacional de Colombia, Ciudad Universitaria, Bogotá, Colombia.*

<sup>2</sup>*Grupo de Identificación Humana e Inmunogenética, Facultad de Medicina, Universidad Nacional de Colombia, Bogotá, Colombia.*

<sup>3</sup>*Instituto de Genética, Servicios Médicos Yunis Turbay y Cia, Bogotá, Colombia.*

<sup>4</sup>*Brigham and Womens Hospital, Department of Pathology, Harvard Medical School, Boston, USA.*

### Abstract

We analyzed 1041 individuals belonging to 17 Amerindian tribes of Colombia, Chimila, Bari and Tunebo (Chibcha linguistic family), Embera, Waunana (Choco linguistic family), Puinave and Nukak (Maku-Puinave linguistic families), Cubeo, Guanano, Tucano, Desano and Piratapuyo (Tukano linguistic family), Guahibo and Guayabero (Guayabero Linguistic Family), Curripaco and Piapoco (Arawak linguistic family) and Yucpa (Karib linguistic family). for MHC class II haplotypes (HLA-DRB1, DQA1, DQB1). Approximately 90% of the MHC class II haplotypes found among these tribes are haplotypes frequently encountered in other Amerindian tribes. Nonetheless, striking differences were observed among Chibcha and non-Chibcha speaking tribes. The DRB1\*04:04, DRB1\*04:11, DRB1\*09:01 carrying haplotypes were frequently found among non-Chibcha speaking tribes, while the DRB1\*04:07 haplotype showed significant frequencies among Chibcha speaking tribes, and only marginal frequencies among non-Chibcha speaking tribes. Our results suggest that the differences in MHC class II haplotype frequency found among Chibcha and non-Chibcha speaking tribes could be due to genetic differentiation in Mesoamerica of the ancestral Amerindian population into Chibcha and non-Chibcha speaking populations before they entered into South America.

**Keywords:** MHC, Class II, Amerindians, Colombia, linguistic.

Received: August 28, 2012; Accepted: January 29, 2013.

### Introduction

The arrival of Native American through Beringia occurred around 12,000-15,000 years before present based on archeological, mtDNA, and Y-chromosome microsatellites haplotypes/SNP studies (Crawford 1988; Salzano and Callegari-Jaques 1988; Santos *et al.*, 1996; Bonatto and Salzano 1997a; 1997b; Santos *et al.*, 1999; Dillehay 2009; O'Rourke and Raff 2010). Recent studies using high resolution SNP genotyping analysis indicate that Native American populations, including all Amerindian populations derived from an Asian ancestral populations with at least two additional streams of gene flow from Asia (Reich *et al.*, 2012).

The population of Colombia is highly diverse (Yunis *et al.*, 2000). The majority of the Colombian contemporary population derived from Spaniards and Amerindians admixture, with a lower contribution of other European (French, Italian, German, and Portuguese) as well as Arabs,

and Jews. These Mestizo populations are located in the Andean, Pacific, Atlantic regions and, in a minor degree, in the Orinoquian and Amazonian regions. On the other hand, populations with a higher proportion of African ancestry, derived from slave trading during the 17th and 18th centuries, are located in the Pacific coast, Caribbean coast and islands. The Amerindian populations, with 81 tribes are located in the plains (Orinoquian region), Amazonian jungle and in some regions of the Colombian Andes. In America, Brazil has the highest number of Amerindian tribes followed by Colombia (81 Tribes), Mexico (68 tribes) and Venezuela (20 tribes) (Salzano and Callegari-Jaques 1988; Yunis *et al.*, 1994, 2001).

Different studies of MHC class-I and class-II alleles in Amerindian populations have been carried out for anthropological as well as evolutionary studies (Layrisse *et al.*, 1973; Kostyu and Amos 1981; Williams *et al.*, 1981; Vullo *et al.*, 1984; Gorodezky *et al.*, 1985; Long *et al.*, 1991; Petzl-Erler *et al.*, 1993; Tsuneto *et al.*, 2003; Parolin and Carnese 2009; Arnaiz-Villena *et al.*, 2010, 2011; Vargas-Alarcon *et al.*, 2011). New MHC class I and class II alleles have been identified in North and South Amerindian

communities (Belich *et al.*, 1992; Watkins *et al.*, 1992; Zhang *et al.*, 1993; Layrisse *et al.*, 1997; Mack and Erlich 1998). The analysis of MHC along with other genetic markers such as mtDNA, Y-chromosome haplotypes and SNP might provide important information regarding the peopling of the continent. Due to its geographical location, Colombia was a necessary stepping-stone for entrance of Amerindian populations into South America. Thus, the study of the actual living Amerindian populations in Colombia should shed some light regarding migrations, admixture, and linguistic relationship.

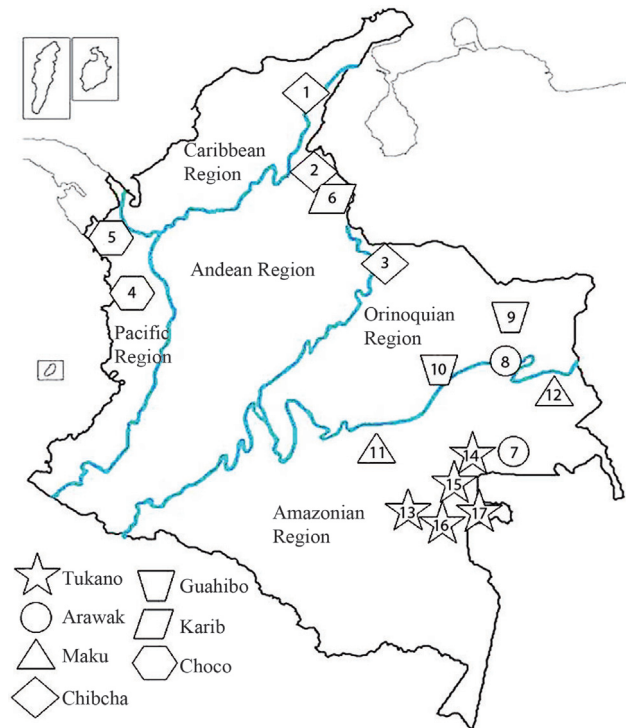
We have previously reported the MHC class II data obtained from four Amerindian tribes located in the northern section of Colombia (Arhuaco, Arsario, Kogui and Wayuu) (Yunis *et al.*, 1994), as well as the MHC class II results obtained from three Amerindian tribes located in the southwest section of Colombia (Guambiano, Paez and Ingano) (Yunis *et al.*, 2001). In these reports we found a correlation between genetics and linguistic affiliation for the Chibcha speaking tribes, demonstrating that the Guambiano and Paez Amerindian tribes (formerly classified as Chibcha speaking populations) are not genetically related to the Chibcha speaking tribes of northern Colombia (Yunis *et al.*, 1994, 2001).

Here, we report the MHC class II haplotype data (DRB1, DQA1 and DQB1) from 17 additional Amerindian tribes located in the Amazonian, Orinoquian, Pacific and Perija mountain range regions of Colombia, plus an analysis with the data from those seven Amerindian tribes previously reported. Significant differences between Chibcha speaking tribes and non-Chibcha speaking tribes were found. These results suggest that the genetic differentiation between Chibcha and non-Chibcha Amerindian groups occurred in Mesoamerica before they entered and spread throughout South America.

## Material and Methods

### Populations studied

We have analyzed 1041 individuals from 17 different Amerindian tribes of Colombia that belong to 7 different linguistic families (Arawak, Chibcha, Choco, Guahibo, Karib, Maku and Tukano) (Figure 1, Table 1). Blood samples were collected between 1989-1992 after proper informed consent was obtained, including the approval of each Chief or Governors of each Tribe. The analysis included the data obtained from the Arhuaco, Kogui, Arsario (Chibcha), Wayuu (Arawak), Ingano (Quechua), Paez and Guambiano (no linguistic classification at present) tribes had been reported previously (Yunis *et al.*, 1994, 2001). Among each tribe there were unrelated individuals and family groups based on the information derived when the samples were obtained. The geographical location and their linguistic affiliation are shown in Table 1. Most of these tribes are in-



**Figure 1** - Geographical location of Amerindian tribes of Colombia (see Table 1). Lines within the country identify the natural regions of Colombia. Different shapes were used to identify the linguistic affiliation of each tribe. Numbers identify each population. Chibcha speaking populations: Chimila (1), Bari (2), Tunebo (3); Choco speaking populations: Wau-nana/Noanama (4) and Embera/Katio (5); Karib speaking population: Yuca/Yuko (6); Arawak speaking populations: Curripaco (7) and Piapoco (8); Guahibo speaking populations: Guahibos/Sikuane (9) and Guayabero (10); Maku speaking populations: Nukak (11) and Puinave (12); Tukano speaking populations: Cubeo (13), Guanano/Wanano (14), Tukano (15), Desano (16), Piratapuyo (17).

habitants of the Orinoquian flats, the Amazonian, the Pacific region and the Perija Mountain range.

### DNA isolation

DNA was isolated from ACD (anticoagulant citrate dextrose) preserved blood by a quick lysis method (Kawasaki 1990), or by a salting-out method with minor modifications (Miller *et al.*, 1988).

### PCR Amplifications

MHC class II typing was carried as described before. The second exon of the DRB1 generic locus, and the DQA1 and DQB1 loci were amplified by PCR from genomic DNA. The primers and conditions used in this study have been published elsewhere (Salazar *et al.*, 1992; Yunis *et al.*, 1992, 1994, 2001). In addition, allele specific amplifications for DRB1\*15/16 and DRB1\*04 were performed as described before (Yunis *et al.*, 1994).

The DRB generic and high resolution (except for DRB1\*04 alleles), DQA1 and DQB1 alleles were determined in the locus-specific PCR amplified products by sequence-specific oligonucleotide probe hybridization

**Table 1** - Number of individuals (N), linguistic family affiliation (based on Ruhlen, 1991) and geographical location of Amerindian Colombian Tribes analyzed.

Tribe	N	Linguistic family	Geographical location	Number in Figure 1
CHIMILA	88	CHIBCHA	Sierra Nevada Santa Marta	1
BARI	90	CHIBCHA	Perija Mountain Range	2
TUNEBO/ U'WA	29	CHIBCHA	Orinoquian region	3
EMBERA/CATIO	73	CHOCO	Pacific region	4
WAUNANA/NOANAMA	60	CHOCO	Pacific region	5
YUCPA	47	KARIB	Perija Mountain Range	6
CURRIPACO	49	ARAWAK	Orinoquian region	7
PIAPOCO	25	ARAWAK	Orinoquian region	8
GUAHIBOS/SIKUANE	34	GUAHIBO	Orinoquian region	9
GUAYABERO	42	GUAHIBO	Orinoquian region	10
PUINAVE	175	MAKU	Amazonian North	11
NUKAK	29	MAKU	Amazonian North	12
CUBEO	72	TUKANO	Amazonian North	13
GUANANO	118	TUKANO	Amazonian North	14
TUCANO	37	TUKANO	Amazonian North	15
DESANO	49	TUKANO	Amazonian North	16
PIRATAPUYO	24	TUKANO	Amazonian North	17
Total	1041			

(PCR-SSO) labeled with alkaline phosphates and chemiluminescent detection (Lifecodes Corporation, Stamford, CT). For DRB1\*04 alleles, a PCR-RFLP based method was used as described before (Yunis *et al.*, 1994).

### Data analysis

Haplotype frequencies were determined by direct counting of haplotypes in unrelated individuals based on well known and strong linkage disequilibrium MHC class II haplotype associations for DRB1, DQA1 and DQB1 alleles, as has been documented previously on multiple populations around the world (Begovich *et al.*, 1992; Imanishi *et al.*, 1992; Clayton *et al.*, 1997). In addition, analysis within families was conducted for MHC class II segregation. Once a haplotype was identified within a family group, it was counted only once if the offspring carried the same haplotype in order to avoid overestimation of MHC frequencies. Each haplotype from non-consanguineous family members was counted once because they would be segregated to the offspring. An unrelated individual was presumed homozygous when typing results showed the presence of only one haplotype. Within families, homozygosity was based on segregation analysis of haplotypes.

### Genetic distances

Dendrograms produced by the neighbor-joining (NJ) method based on Nei genetic distances (Nei 1972) were calculated from MHC class II haplotype frequencies obtained from the Amerindian tribes analyzed here with the addition of those published before (Yunis *et al.*, 1994,

2001) with the aid of PHYLIP (Felsenstein 1993). Briefly, bootstrapping was used to generate multiple sets (100 sets) from the MHC class II haplotype data. The bootstrapping outfile was then used to calculate Nei genetic distances, followed by the Neighbor-Joining module to generate the NJ dendrograms. Finally, the Consense program was used to obtain a Consense tree based on the NJ tree. A tree was plotted with the aid of TreeView (V32).

### Results

Table 2 describes the MHC class II haplotypes found in seventeen Colombian Amerindian tribes of the Orinoquian, Amazonian and Perija Mountain regions. A limited number of haplotypes accounted for nearly 90% of all HLA class II haplotypes found in all Amerindian populations tested. These haplotypes are frequently found in other Amerindian tribes. However, striking differences were found between populations belonging to different linguistic families in particular, those belonging to the Chibcha and Non-Chibcha speaking groups (Ecuatorial-Tucanoan linguistic family) (Ruhlen 1991).

The DRB1\*04:04, DQA1\*03:01, DQB1\*03:02 haplotype was found in 11 out of 17 tribes with the highest frequency in the Piratapuyo tribe (20.6%), followed by the Tucano, Desano and Nukak tribes. However, the highest frequency reported for this haplotype in Colombian Amerindians was found in the Ingano tribe with a 24.3% frequency, followed by the Guambiano tribe (21%) (Yunis *et al.*, 2001). Of interest, none of the six Chibcha speaking

**Table 2** - MHC class II haplotype (HLA-DRB, DQB1, DQA1) frequencies of Amerindian tribes of Colombia.

DRB1	DRB3	DRB4	DRB5	DQA1	DQB1	NUKAK N = 57	PUNAVE N = 281	GUAYABERO N = 71	PIAPOCO N = 35	CURUPACO N = 77	CUBEO N = 129	GUANANO N = 33	TUCANO N = 58	DESANO N = 80	PIRATAPUYO N = 34	TUNEBO N = 33	CHIMILA N = 113	BARI N = 60	YUCPAS N = 37	EMBERA N = 85	WAINANA N = 79
04:03		01:00		03:01	03:02	0.035	0.043			0.013	0.008				0.029				0.162	0.024	0.038
04:04		01:00		03:01	03:02	0.158	0.053			0.091	0.147				0.206					0.012	
04:07		01:00		03:01	03:02		0.004	0.019			0.008					0.121	0.841	0.283			0.129
04:11		01:00		03:01	03:02	0.281	0.231	0.019	0.086	0.143	0.248	0.182	0.207	0.125	0.088			0.250	0.514		0.094
16:02			02:00	05:01	03:01	0.140	0.068	0.462	0.286	0.182	0.194	0.121	0.138	0.200	0.235	0.303		0.333	0.162		0.282
14:02	01:01			05:01	03:01	0.351	0.107	0.404	0.171	0.195	0.186	0.121	0.069	0.175	0.088	0.333	0.124	0.083	0.135		0.247
14:06	01:01			05:01	03:01		0.007						0.017		0.059						
08:02				04:01	04:02	0.035	0.071	0.019	0.029	0.065	0.070	0.061	0.017	0.063	0.059	0.091	0.018	0.017	0.027	0.059	0.139
08:04		01:00		04:01	04:02		0.292		0.229	0.221	0.054	0.242	0.172	0.075	0.147	0.030					
09:01		01:00		03:01	03:03			0.019			0.047	0.091	0.052	0.063							
04:11		01:00		03:01	04:02		0.007	0.019		0.052	0.008		0.069	0.038	0.029	0.121					
04:17		01:00		03:01	03:02		0.075														
04:1C		01:00		03:01	03:02																
04:1C		01:00		03:01	03:02																
04:02		01:00		03:01	03:02																
04:04		01:00		03:01	03:01																
04:05		01:00		03:01	03:02		0.004														
04:07		01:00		03:01	03:01																
04:08		01:00		03:01	03:02																
04:11		01:00		05:01	03:01				0.029												
04:10		01:00		03:01	04:02		0.007														
04:10		01:00		03:01	03:02								0.017	0.013							
04:10		01:00		03:01	03:02																
04:03/04:10		01:00		ND	03:01		0.004				0.030										
11:01	02:00			05:01	03:01		0.004													0.012	
11:02	02:00			05:01	03:01																
12:01	02:00			05:01	03:01					0.013										0.013	0.025
12:01	02:00			01:01	05:01																
14:02	01:01			04:01	04:02															0.012	
08:00				ND	03:01								0.017	0.025	0.029					0.012	0.038
08:02				04:01	03:02									0.013							
08:04				ND	03:02								0.034								
08:04				05:01	04:02																
03:01	01:01			05:01	02:01		0.007														
03:02	01:01			04:01	04:02		0.004														
03:01	01:01			05:01	03:01																0.013
13:01	01:01			01:03	06:03		0.004														0.013
13:03	01:01			05:01	03:01																
13:04	01:01			05:01	03:01																
01:03				01:01	05:01		0.004		0.029		0.008		0.017								
09:01		01:00		02:01	02:01										0.029						
09:01		01:00		05:01	03:01															0.012	
16:02			02:00	04:01	04:02																
16:02			02:00	03:01	03:01										0.025						
15:03			01:01	05:01	03:01					0.013											
07:01	01:00			02:01	02:01					0.013								0.017			

ND: Not done. N = Number of informative haplotypes. 04:1C: 04:03/04:04/04:07/04:08/04:00 No allele typing was carried out for these possible alleles.



groups including those previously reported by us carry this haplotype (Arhuaco, Kogui, Arsario, Chimila, Bari, Tunebo) and had marginal frequencies in the Choco speaking groups (Embera and Waunana).

On the contrary, the HLA-DRB1\*04:07, DQA1\*03:01, DQB1\*03:02 haplotype was most frequently found among three Chibcha speaking groups Chimila (84.1%), Bari (28.3%) and Tunebo (12.1%) as well as in the members of the Choco linguistic family (Embera and Waunana). A low frequency of this haplotype was found in Amerindian tribes located in the Orinoquian and Amazonian regions of the country. Previous results showed a high frequency of this haplotype among the Chibcha speaking groups of Northern Colombia, Arsario 45% and Kogui 43% (Yunis *et al.*, 1994). This haplotype was also found at high frequencies among the Paez (nearly 30%), Guambianos (21%) and Ingano tribes (13%) (Yunis *et al.*, 2001).

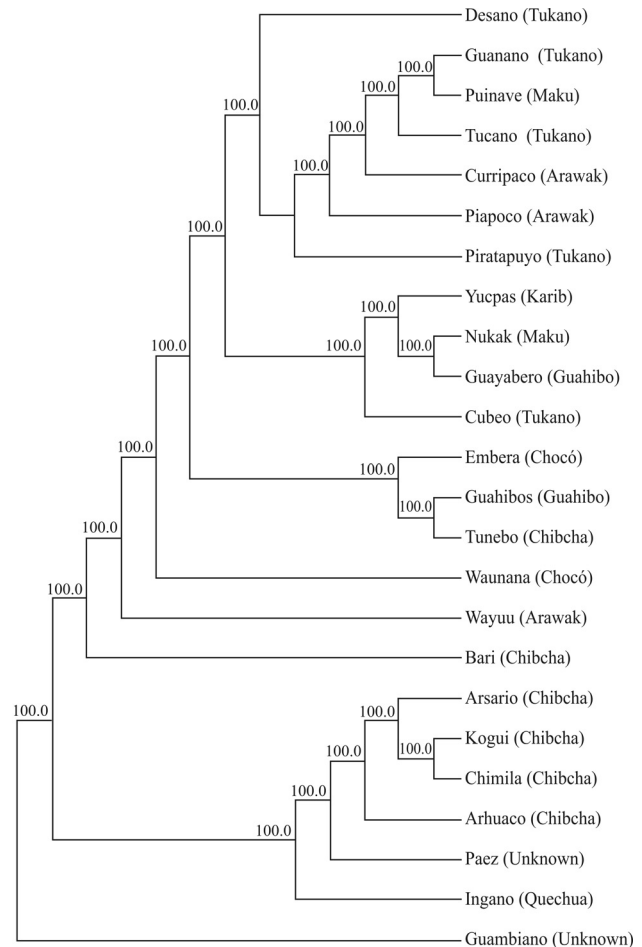
The DRB1\*04:11, DQA1\*03:01, DQB1\*03:02 was found at a 51.4% frequency among the Yucpa (a Karib speaking group) as well as in all Amerindian tribes of the Orinoquian and Amazonian regions. Among the Chibcha speaking groups, this haplotype was only found among the Bari Indians, a tribe that shares the same geographical location with the Yucpa in the Perija mountain range between Colombia and Venezuela. No other Chibcha speaking group carried this haplotype that was found in almost all other Amerindian tribes of the Orinoquian and Amazonian regions of Colombia with frequencies ranging from 2-28%.

The DRB1\*16:02, DQA1\*05:01, DQB1\*03:01 was found in the majority of the Amerindian tribes analyzed with the highest frequency among the Guahibos (46.2%). This haplotype was not found in the Chibcha speaking groups with the exception of the Bari (33.3%) and Tunebo (30.3%) tribes.

The DRB1\*14:02, DQA1\*05:01, DQB1\*03:01 haplotype, is common among all Amerindian tribes. The highest frequency was found in the Guayabero tribe (62%), followed by the Guahibo tribe (40.4%), both of them belonging to the Guahibo linguistic Family.

Two DRB1\*09:01 carrying haplotypes were found in several Amerindian Populations of the Orinoquian region of Colombia. The DRB1\*09:01, DQA1\*03:01, DQB1\*03:03 had the highest frequency among the Tunebo (12.1%) followed by the Tucano, Desano, Piratapuyo, Guahibo, Cubeo and Puinave. The second haplotype, HLA-DRB1\*09:01, DQA1\*03:01, DQB1\*03:02 was found in the Guanano (9.1%), Desano, Tucano, Cubeo, Guahibo and Guayabero tribes. None of the Chibcha speaking groups carried any of the DRB1\*09:01 haplotypes with the exception of the Tunebo tribe that showed the highest frequency for the DRB1\*09:01, DQA1\*03:01, DQB1\*03:03 haplotype (12.1%) among all populations tested.

A neighbor-joining tree was generated as described (Figure 2). Two clusters were identified. In one of them, the Chibcha speaking tribes of Northern Colombia (Arhuaco,



**Figure 2** - Consensus tree based (100 sets) on Nei genetic distances (Nei, 1972) from MHC Class II haplotypes of 24 Amerindian tribes of Colombia. Data for the Arhuaco, Kogui, Arsario, Wayuu, Paez, Ingano, and Guambiano had been published previously (Yunis *et al.*, 1994, 2001). The linguistic family for each tribe is presented in parentheses.

Kogui, Chimila, Arsario) were closely grouped. In that cluster, but distantly related, are also the Paez and Ingano tribes (the latter a Quechua speaking group). The second cluster includes all the Amerindian populations of the Orinoquian/Amazonian region of Colombia analyzed herein. Within this cluster are also included two Amerindian populations belonging to the Choco-Paezan branch of the Chibcha family, Embera and Waunana (Ruhlen 1991). The close proximity of the Embera, Tunebo and Guahibo tribes is probably due to the high frequency of the DRB1\*14:02, DQA1\*05:01, DQB1\*03:01 and DRB1\*16:02, DQA1\*05:01, DQB1\*03:01 haplotypes found in these tribes. The Bari, another Chibcha speaking tribe, clustered within this group, but is more distantly related. This result is probably due to gene flow between the Yucpa tribe, a Karib Speaking group, as has been previously documented.

## Discussion

Different studies have been carried out in order to analyze the genetic variability and its evolutionary implications for

Amerindian tribes. Among them, mtDNA, Y-chromosome and different autosomic including the MHC markers have been used. A recent review of uniparental genetic markers in South Amerindians was published (Bisso-Machado *et al.*, 2012). We have performed MHC class II analysis in a large sample of Amerindian individuals belonging to different linguistic families. Our results have shown marked differences among members of different linguistic families. In particular Chibcha speaking and non-Chibcha speaking groups showed contrasting differences, especially for DRB1\*04:04; DRB1\*04:07; DRB1\*04:11 and DRB1\*09:01 carrying haplotypes.

The DRB1\*04:04, DRB4\*01, DQA1\*03:01, DQB1\*03:02 haplotype had originally been described in European-descendent populations (Fernandez-Vina *et al.*, 1991). However, this haplotype has also been found in Amerindians of the southwestern region of North America (4.4%) (Miller *et al.*, 1992), and at significant frequencies in some isolated South Amerindian tribes of Colombia, as well as, in the Cayapa Indians of Ecuador, the Kaingang, Guarani and Xavante tribes of Brazil, and in the Toba and Mataco-Wichi tribes of Argentina (Petzl-Erler *et al.*, 1993, 1997; Titus-Trachtenberg *et al.*, 1994; Trachtenberg *et al.*, 1995, 1996; Tsuneto *et al.*, 2003). In addition, this haplotype had marginal frequencies in one tribe belonging to the Choco linguistic family (Embera) and in one Guahibo speaking tribe (Guayabero). Of interest, this haplotype was not found in any of the six Chibcha speaking Amerindian tribes of Colombia previously analyzed by us (Yunis *et al.*, 1994) nor in the Chimila, Bari, and Tunebo tribes analyzed in the present report. Likewise, it was not present in the Bari and Warao Indians of Venezuela (Chibcha linguistic family) (Guedez *et al.*, 1994; Petzl-Erler *et al.*, 1997). The only exception for the presence of the DRB1\*04:04, DRB4\*01, DQA1\*03:01, DQB1\*03:02 haplotype in a Chibcha speaking tribe was among the Cayapa Indians of Ecuador, as reported by others (Titus-Trachtenberg *et al.*, 1994; Trachtenberg *et al.*, 1995). In this regard, some authors have postulated that the Cayapa Indians originated in the Amazonian region and then migrated to the Andes and later to the coastal region of Ecuador. Other researchers have reported that the Cayapa tribe originated in the Andean highlands in the northern area of Ecuador, and as a result of the expansion of the Inca empire during the 15th century and the Spanish invasion in the 16th, they moved toward the coast of Ecuador (Barret 1925; Barriga Lopez 1987; Carrasco 1988; Stinson 1989). The presence of the DRB1\*04:04, DRB4\*01, DQA1\*03:01, DQB1\*03:02 haplotype in the Cayapa Indians could be explained by gene flow from other Amerindian tribes, in particular from the Inca population. In this regard, The Ingano tribe located in the southwest section of Colombia, a Quechua speaking group, are direct descendants of the Inca empire and showed the highest frequency of the DRB1\*04:04, DRB4\*01, DQA1\*03:01, DQB1\*03:02 haplotype (24.3%)

in our previous study (Yunis *et al.*, 2001). Thus, it is possible to think that gene flow had occurred that accounts for the presence of this haplotype among the Cayapa Indians. The presence of The DRB1\*04:04, DRB4\*01, DQA1\*03:01, DQB1\*03:02 haplotype among Amerindians of North, Central and South America as well as among European descent populations indicates that this haplotype represents an ancient HLA haplotype.

The DRB1\*04:07, DQA1\*03:01, DQB1\*03:02 is one of the most frequent MHC class II haplotypes found among Chibcha speaking groups. The highest frequency found was among the Chimila tribe (84.1%), followed by the Arsario, Kogui, Bari, Arhuaco and Tunebo, as well as in the Choco-Paezan speaking groups Embera and Waunana of the Chibcha family. This haplotype had also been found in Central America among the Mexican Amerindian tribes, Seri (44%), Mazatecan (28%), and Mixteco (29%) (Arnaiz-Villena *et al.*, 2000; Salzano 2002), among the Cayapa Amerindian tribe of Ecuador (27%), the Wayuu tribe of Colombia (11.5%), the Guarani in Brazil (Tsuneto *et al.*, 2003), and in the Mapuche and Tehuelche tribes of Argentina (Parolin and Carnese 2009). In contrast, it was present at very low frequencies among few Amerindian tribes of the Orinoquian and Amazonian region of Colombia (Guahibo 1.9%, Cubeo 0.8% and Puinave 0.4%).

The DRB1\*04:11, DQA1\*03:01, DQB1\*03:02 haplotype is a common haplotype among Amerindian tribes of the Orinoquian region of Colombia with frequencies ranging from 2% (Guahibo) to 28.1% (Nukak). Among Chibcha speaking groups, only the Bari tribe carried this haplotype with a frequency of 25%. The Yucpa tribe (a Karib speaking group) shares the same geographical location with the Bari tribe in the Perija mountain range and showed the highest frequency for this haplotype (51.4%). In a previous report, this DRB1\*04:11 haplotype had been described at a 60% frequency among the Yucpa Amerindian tribe of Venezuela (Layrisse *et al.*, 2001) and among the Aché in Brazil (74.1%) (Tsuneto *et al.*, 2003) as the predominant MHC class II haplotype. In addition, this haplotype is present in the Wayuu tribe (Arawak linguistic family) located in the north section of Colombia (Yunis *et al.*, 1994). The presence of Arawak and Karib speaking groups in the northern section of Colombia and Venezuela is the result of the expansion/migration of these tribes from the Amazonian region towards the north part of South America that later populated the Caribbean islands. Thus, the presence DRB1\*04:11 carrying haplotype among the Bari tribe could be explained by gene flow from those Amerindian tribes that migrated from the Amazonian region towards the north section of Colombia.

It is of interest that the DRB1\*09:01, DQA1\*03:01, DQB1\*03:03 haplotype showed the highest frequency in the Tunebo tribe (Chibcha linguistic family). This haplotype was not found in any other Chibcha speaking tribe of Colombia or Venezuela. It was found at frequencies be-

tween near 1% and 7% in some Amerindian tribes of the Orinoquian and Amazonian Regions (Puinave, Guahibo, Cubeo, Tucano, Desano, and Piratapuyo). From all these tribes, only the Guahibo is closer geographically to the Tunebo. However, the frequency of this haplotype among the Guahibo was only 1.9% and a similar frequency (2%) for the Guahibo (also known as Sikuni) had been reported previously (Trachtenberg *et al.*, 1996). Thus, the high frequency of this haplotype among the Tunebo remains to be explained.

The NJ tree generated based on MHC class II frequencies (Figure 2) showed mainly the presence of two clusters. In one of them, the Chibcha speaking tribes of Northern Colombia (Arhuaco, Kogui, Chimila, Arsario). More distantly within that cluster are the Ingano (Quechua) and Paez (language without classification at present). The other cluster includes all the Amerindian tribes of the Orinoquian and Amazonian region, as well as the Embera and Waunana (Choco speaking groups) of the Pacific region of Colombia near Panama. In this regard, it is worth mentioning that previous studies on mtDNA and nuclear genetic markers have shown significant differences between the Chibcha speaking groups of Panama (Kuna and Ngobe) and the Embera and Waunana tribes (Kolman and Bermingham 1997) of the Choco-Paezan branch of the Chibcha macrophilum language. These results are relevant, since recent work based on SNP analysis shows that the Embera and Waunana tribes are genetically distant from the Chibcha speaking tribes of Colombia and Central America (Reich *et al.*, 2012), results that are in agreement with our results based on MHC class II (Figure 2).

The presence of some MHC class II haplotypes frequently found among Amerindian tribes of the Orinoquian and Amazonian regions (Equatorial-Tucano and Ge-Pano-Carib linguistic families) among few Chibcha speaking tribes could be explained due to the geographical expansion and gene flow of the former tribes towards their present location. An example of that expansion is the Barí tribe in the Perija Mountain range between Colombia and Venezuela where gene flow has been documented with the Yuca Tribe (Karib speaking) (Layrisse *et al.*, 1997, 2001).

Previous studies have shown a relationship between Chibcha speaking tribes of Central America and Chibcha speaking tribes of Northern Colombia based not only on genetic data like mtDNA (Melton *et al.*, 2007) but also based on cultural aspects like shared settlement patterns, iconography, material goods and linguistics (Lange and Stone 1984; Hoopes and Fonseca 2003).

The differences found between the Chibcha speaking groups of Colombia and non-Chibcha Amerindian populations for MHC class II haplotypes suggest that the Chibcha speaking populations could have originated from the ancestral population in Mesoamerica and then differentiated from the rest of the Amerindian linguistics branches before spreading southwards to South America. Additional studies

carried out by us and others, analyzing mtDNA haplogroups (Keyeux *et al.*, 2002; Melton *et al.*, 2007) have also shown differences between Chibcha and non-Chibcha speaking populations (unpublished data).

Recently, it has been postulated that reverse gene flow between Chibcha speaking tribes located at the Isthmus-Colombian area could have occurred across the Panama isthmus after the initial settlement in South America, in particular, based on the genetic findings of the Cabecar tribe of Costa Rica (Reich *et al.*, 2012). Although this is a plausible conclusion, that study only included two Chibcha speaking populations (Kogui and Arhuaco) and two populations belonging to the Paezan-Choco branch of the Chibcha linguistic family (Embera and Waunana) from Colombia. The same authors also indicated an alternative scenario whereby the Chibcha ancestral populations differentiated from the rest of the Amerindian population before spreading in Central and South America. Additional studies including many other Chibcha speaking tribes such as Arsario, Tunebo, Barí, Warao and Chimila from Colombia, as well as Chibcha Amerindian tribes of Panama such as the Ngobe and Kuna, among others, will help to shed some light in that regard. In addition, the simultaneous analysis of mtDNA, Y-chromosome SNP/microsatellite and autosomic markers of Amerindian tribes from North, Central and South America should provide important evidence that will eventually lead to a better understanding of the differences detected in the present study.

## Acknowledgments

We would like to especially thank the members of all the Amerindian tribes of Colombia and their leaders that kindly and voluntarily donated samples for this study. This work was supported in part by a grant from COLCIENCIAS # 1101-05-002-9 to Emilio J. Yunis.

## References

- Arnaiz-Villena A, Vargas-Alarcon G, Granados J, Gomez-Casado E, Longas J, Gonzales-Hevilla M, Zuniga J, Salgado N, Hernandez-Pacheco G, Guillen J, *et al.* (2000) HLA genes in Mexican Mazatecs, the peopling of the Americas and the uniqueness of Amerindians. *Tissue Antigens* 56:405-416.
- Arnaiz-Villena A, Parga-Lozano C, Moreno E, Areces C, Rey D and Gomez-Prieto P (2010) The origin of Amerindians and the peopling of the Americas according to HLA genes: Admixture with Asian and Pacific people. *Curr Genomics* 11:103-114.
- Arnaiz-Villena A, Parga-Lozano C, Moreno E, Areces C, Gomez-Prieto P, de Salamanca ME and Rey D (2011) The origin of Amerindians and the peopling of the Americas according to HLA genes: Admixture with Asian and Pacific people. *Tissue Antigens* 77:491-491.
- Barret SA (1925) The Cayapa Indians of Ecuador. *Indian Notes and Monographs*. Heye Foundation, New York, 476 pp.



- Barriga Lopez F (1987) Cayapas o Chiachis. Vol 4: Etnología Ecuatoriana. Gráficas Duque, Quito, 320 pp.
- Begovich AB, McClure GR, Suraj VC, Helmuth RC, Fildes N, Bugawan TL, Erlich HA and W. K (1992) Polymorphism, recombination, and linkage disequilibrium within the HLA class II region. *J Immunol* 1:249-258.
- Belich MP, Madrigal JA, Hildebrand WH, Zemmour J, Williams RC, Luz R, Petzl-Erler ML and Parham P (1992) Unusual HLA-B alleles in two tribes of Brazilian Indians. *Nature* 357:326-329.
- Bisso-Machado R, Bortolini MC and Salzano FM (2012) Uniparental genetic markers in South Amerindians. *Genet Mol Biol* 35:365-387.
- Bonato SL and Salzano FM (1997a) A single and early migration for the peopling of the Americas supported by mitochondrial DNA sequence data. *Proc Natl Acad Sci USA* 94:1866-1871.
- Bonato SL and Salzano FM (1997b) Diversity and age of the four major mtDNA haplogroups, and their implications for the peopling of the New World. *Am J Hum Genet* 61:1413-1423.
- Carrasco E (1988) El Pueblo Chachi: El Jeengume Avanza. Abya-Yala Ediciones, Quito, 216 pp.
- Clayton J, Lonjou C and D. W (1997) Allele and haplotype frequencies for HLA loci in various ethnic groups. HLA, Genetic Diversity of HLA Functional and Medical Implication. Proceedings of the Twelfth International Histocompatibility Workshop and Conference, EDK Medical and Scientific International Publisher, Saint Malo, France, 867 pp.
- Crawford MH (1988) The Origin of Native Americans. Evidence from Anthropological Genetics. Cambridge University Press, Cambridge, 326 pp.
- Dillehay TD (2009) Probing deeper into first American studies. *Proc Natl Acad Sci USA* 106:971-978.
- Felsenstein J (1993) PHYLIP (Phylogeny Inference Package) ver. 3.5c, distributed by the author. Department of Genetics, University of Washington, Seattle.
- Fernandez-Vina MA, Gao XJ, Moraes ME, Moraes JR, Salatiel I, Miller S, Tsai J, Sun YP, An JB, Layrisse Z, *et al.* (1991) Alleles at four HLA class II loci determined by oligonucleotide hybridization and their associations in five ethnic groups. *Immunogenetics* 34:299-312.
- Gorodezky C, Castro-Escobar LE and Escobar-Gutierrez A (1985) The HLA system in the prevalent Mexican Indian group: The Nahuas. *Tissue Antigens* 25:38-46.
- Guedez YB, Layrisse Z, Dominguez E, Rodriguez-Larralde A and Scorza J (1994) Molecular analysis of MHC class II alleles and haplotypes (DRB1, DQA1 and DQB1) in the Bari Amerindians. *Tissue Antigens* 44:125-128.
- Hoopes J and Fonseca O (2003) Goldwork and Chibchan Identity: Endogenous Change and Diffuse Unity in the Isthmo-Colombian Area. Gold and Power in Ancient Costa Rica, Panama and Colombia. Dumbarton Oaks, Washington DC, 90 pp.
- Imanishi T, Akaza T, Kimura A, Tokunaga K and TG (1992) Allele and haplotype frequencies for HLA and complement loci in various ethnic groups. In: HLA 1991: Proceedings of the Eleventh Histocompatibility Workshop and Conference. Oxford Science Publications, Yokohama, 1220 pp.
- Kawasaki ES (1990) Sample preparation from Blood, Cells and other fluids. In: PCR Protocols: A Guide to Methods and Applications. Academic Press Inc., San Diego, 482 pp.
- Keyeux G, Rodas C, Gelvez N and Carter D (2002) Possible migration routes into South America deduced from mitochondrial DNA studies in Colombian Amerindian populations. *Am J Hum Genet* 74:211-233.
- Kolman CJ and Bermingham E (1997) Mitochondrial and nuclear DNA diversity in the Choco and Chibcha Amerinds of Panama. *Genetics* 147:1289-1302.
- Kostyu DD and Amos DB (1981) Mysteries of the Amerindians. *Tissue Antigens* 17:111-123.
- Lange FW and Stone DZ (1984) The Archeology of Lower Central America. SAR Press, Albuquerque, 490 pp.
- Layrisse Z, Layrisse M, Malave I, Terasaki P, Ward RH and Neel JV (1973) Histocompatibility Antigens in a genetically isolated American indian tribe. *Am J Hum Genet* 25:493-509.
- Layrisse Z, Matos M, Montagnani S and Herrera F (1997) The novel HLA-DRB1\*0807 allele is present in a Carib Amerindian population: The Yucpa. *Tissue Antigens* 49:427-428.
- Layrisse Z, Guedez Y, Dominguez E, Paz N, Montagnani S, Matos M, Herrera F, Ogando V, Balbas O and Rodriguez-Larralde A (2001) Extended HLA haplotypes in a Carib Amerindian population: The Yucpa of the Perija Range. *Hum Immunol* 62:992-1000.
- Long JC, Williams RC, McAuley JE, Medis R, Partel R, Tregellas WM, South SF, Rea AE, McCormick SB and Iwaniec U (1991) Genetic variation in Arizona Mexican Americans: estimation and interpretation of admixture proportions. *Am J Phys Anthropol* 84:141-157.
- Mack SJ and Erlich HA (1998) HLA class II polymorphism in the Ticuna of Brazil: Evolutionary implications of the DRB1\*0807 allele. *Tissue Antigens* 51:41-50.
- Melton PE, Briceño I, Gómez A, Devor EJ, Bernal JE and Crawford MH (2007) Biological relationship between Central and South American Chibchan speaking populations: Evidence from mtDNA. *Am J Phys Anthropol* 133:753-770.
- Miller SA, Dykes DD and Polesky HF (1988) A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res* 16:1215.
- Miller SN, Fernandez-Viña M and Stastny P (1992) Serology and DNA typing for HLA class II in North American Indians and Blacks. In: HLA 1991 Proceedings of the Eleventh International Histocompatibility workshop and conference. Oxford Science Publications, Yokohama, 1220 pp.
- Nei M (1972) Genetic distance between populations. *Am Nat* 106:283-292.
- O'Rourke DH and Raff JA (2010) The human genetic history of the Americas: The final frontier. *Curr Biol* 20:R202-207.
- Parolin ML and Carnese FR (2009) HLA-DRB1 alleles in four Amerindian populations from Argentina and Paraguay. *Genet Mol Biol* 32:212-219.
- Petzl-Erler ML, Luz R and Sotomaior VS (1993) The HLA polymorphism of two distinctive South-American Indian tribes: The Kaingang and the Guarani. *Tissue Antigens* 41:227-237.
- Petzl-Erler ML, Gorodezky C, Layrisse Z, Klitz W, Fainboim L, Vullo C, Bodmer JG, Egea E, Navarrete C, Infante E, *et al.* (1997) Anthropology report for region Latin-America: Amerindian and admixed populations. In: Proceedings of the Twelfth International Histocompatibility Workshop and

- Conference. Genetic Diversity of HLA. Functional and Medical Implications. Medical and Scientific International Publisher, Paris, 867 pp.
- Reich D, Patterson N, Campbell D, Tandon A, Mazieres S, Ray N, Parra MV, Rojas W, Duque C, Mesa N, *et al.* (2012) Reconstructing Native American population history. *Nature* 488:370-374.
- Ruhlen M (1991) *A Guide to the World's Languages. Volume 1: Classification.* Stanford University Press, Stanford, California, 463 pp.
- Salazar M, Yunis JJ, Delgado MB, Bing D and Yunis EJ (1992) HLA-DQB1 allele typing by a new PCR-RFLP method: Correlation with a PCR-SSO method. *Tissue Antigens* 40:116-123.
- Salzano FM and Callegari-Jaques SM (1988) South American Indians. A Case Study in Evolution. Clarendon Press, Oxford, 259 pp.
- Salzano FM (2002) Molecular variability in Amerindians: Widespread but uneven information. *An Acad Bras Cienc* 74:223-263.
- Santos FR, Rodriguez-Delfin L, Pena SD, Moore J and Weiss KM (1996) North and South Amerindians may have the same major founder Y chromosome haplotype. *Am J Hum Genet* 58:1369-1370.
- Santos FR, Pandya A, Tyler-Smith C, Pena SD, Schanfield M, Leonard WR, Osipova L, Crawford MH and Mitchell RJ (1999) The central Siberian origin for native American Y chromosomes. *Am J Hum Genet* 64:619-628.
- Stinson S (1989) Physical growth of Ecuadorian Chachi Amerindians. *Am J Hum Biol* 1:697-707.
- Titus-Trachtenberg EA, Rickards O, De Stefano GF and Erlich HA (1994) Analysis of HLA class II haplotypes in the Cayapa Indians of Ecuador: A novel DRB1 allele reveals evidence for convergent evolution and balancing selection at position 86. *Am J Hum Genet* 55:160-167.
- Trachtenberg EA, Erlich HA, Rickards O, DeStefano GF and Klitz W (1995) HLA class II linkage disequilibrium and haplotype evolution in the Cayapa Indians of Ecuador. *Am J Hum Genet* 57:415-424.
- Trachtenberg EA, Keyeux G, Bernal JE, Rhodas MC and Erlich HA (1996) Results of Expedición Humana. I. Analysis of HLA class II (DRB1-DQA1-DPB1) alleles and DR-DQ haplotypes in nine Amerindian populations from Colombia. *Tissue Antigens* 48:174-181.
- Tsuneto LT, Probst CM, Hutz MH, Salzano FM, Rodriguez-Delfin LA, Zago MA, Hill K, Hurtado AM, Ribeiro-dos-Santos AK and Petzl-Erler ML (2003) HLA class II diversity in seven Amerindian populations. Clues about the origins of the Ache. *Tissue Antigens* 62:512-526.
- Vargas-Alarcon G, Granados J, Perez-Hernandez N, Rodriguez-Perez JM, Canto-Cetina T, Coral-Vazquez RM, Areces C, Gomez-Prieto P and Arnaiz-Villena A (2011) HLA-class II genes in Mexican Amerindian Mayas: Relatedness with Guatemalan Mayans and other populations. *Immunol Invest* 40:101-111.
- Vullo CM, Celis EM, Serra HM and Riera CM (1984) Study of HLA system in a Mataco population: A geographically isolated American Indian tribe. *Tissue Antigens* 23:33-40.
- Watkins DI, McAdam SN, Liu X, Strang CR, Milford EL, Levine CG, Garber TL, Dogon AL, Lord CI, Ghim SH, *et al.* (1992) New recombinant HLA-B alleles in a tribe of South American Amerindians indicate rapid evolution of MHC class I loci. *Nature* 357:329-333.
- Williams RC, Morse HG, Bonnell MD, Rate RG and Kuberski TT (1981) The HLA loci of the Hopi and Navajo. *Am J Phys Anthropol* 56:291-296.
- Yunis JJ, Delgado MB, Lee-Lewandroski E, Yunis EJ and Bing DH (1992) Rapid identification of HLA-DRw53-positive samples by a generic DRB-PCR amplification without further analysis. *Tissue Antigens* 40:41-44.
- Yunis JJ, Ossa H, Salazar M, Delgado MB, Deulofeut R, de la Hoz A, Bing DH, Ramos O and Yunis EJ (1994) Major histocompatibility complex class II alleles and haplotypes and blood groups of four Amerindian tribes of northern Colombia. *Hum Immunol* 41:248-258.
- Yunis JJ, Garcia O, Baena A, Arboleda G, Uriarte I and Yunis E (2000) Population frequency for the short tandem repeat loci D18S849, D3S1744, and D12S1090 in Caucasian-Mestizo and African descent populations of Colombia. *J Forensic Sci* 45:429-431.
- Yunis JJ, Yunis EJ and Yunis E (2001) Genetic relationship of the Guambino, Paez, and Ingano Amerindians of southwest Colombia using major histocompatibility complex class II haplotypes and blood groups. *Hum Immunol* 62:970-978.
- Zhang S, Fernandez-Vina M, Falco M, Cerna M, Raimondi E and Stastny P (1993) A novel HLA-DRB1 allele (DRB1\*0417) in South American Indians. *Immunogenetics* 38:463.

*Associate Editor: Mara H. Hutz*

License information: This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.